



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Development of a chicken 5 K microarray targeted towards immune function

**Citation for published version:**

Smith, J, Speed, D, Hocking, PM, Talbot, RT, Degen, WGJ, Schijns, VEJC, Glass, EJ & Burt, DW 2006, 'Development of a chicken 5 K microarray targeted towards immune function' BMC Genomics, vol 7, 49, pp. 49. DOI: 10.1186/1471-2164-7-49

**Digital Object Identifier (DOI):**

[10.1186/1471-2164-7-49](https://doi.org/10.1186/1471-2164-7-49)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

BMC Genomics

**Publisher Rights Statement:**

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Methodology article

Open Access

## Development of a chicken 5 K microarray targeted towards immune function

Jacqueline Smith\*<sup>1</sup>, David Speed<sup>1</sup>, Paul M Hocking<sup>1</sup>, Richard T Talbot<sup>2</sup>, Winfried GJ Degen<sup>3</sup>, Virgil EJC Schijns<sup>3</sup>, Elizabeth J Glass<sup>1</sup> and David W Burt<sup>1</sup>

Address: <sup>1</sup>Division of Genetics and Genomics, Roslin Institute, Roslin (Edinburgh), Midlothian, EH25 9PS, UK, <sup>2</sup>Ark-Genomics, Roslin Institute, Roslin (Edinburgh), Midlothian, EH25 9PS, UK and <sup>3</sup>Intervet International B.V., Dept. of Vaccine Technology and Immunology R&D, P.O. Box 31, 5830 AA, Boxmeer, The Netherlands

Email: Jacqueline Smith\* - [Jacqueline.smith@bbsrc.ac.uk](mailto:Jacqueline.smith@bbsrc.ac.uk); David Speed - [David.speed@bbsrc.ac.uk](mailto:David.speed@bbsrc.ac.uk);

Paul M Hocking - [Paul.hocking@bbsrc.ac.uk](mailto:Paul.hocking@bbsrc.ac.uk); Richard T Talbot - [Richard.Talbot@bbsrc.ac.uk](mailto:Richard.Talbot@bbsrc.ac.uk);

Winfried GJ Degen - [Winfried.Degen@intervet.com](mailto:Winfried.Degen@intervet.com); Virgil EJC Schijns - [Virgil.Schijns@intervet.com](mailto:Virgil.Schijns@intervet.com); Elizabeth J Glass - [Liz.glass@bbsrc.ac.uk](mailto:Liz.glass@bbsrc.ac.uk);

David W Burt - [Dave.burt@bbsrc.ac.uk](mailto:Dave.burt@bbsrc.ac.uk)

\* Corresponding author

Published: 13 March 2006

Received: 17 November 2005

BMC Genomics 2006, 7:49 doi:10.1186/1471-2164-7-49

Accepted: 13 March 2006

This article is available from: <http://www.biomedcentral.com/1471-2164/7/49>

© 2006 Smith et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The development of microarray resources for the chicken is an important step in being able to profile gene expression changes occurring in birds in response to different challenges and stimuli. The creation of an immune-related array is highly valuable in determining the host immune response in relation to infection with a wide variety of bacterial and viral diseases.

**Results:** Here we report the development of chicken immune-related cDNA libraries and the subsequent construction of a microarray containing 5190 elements (in duplicate). Clones on the array originate from tissues known to contain high levels of cells related to the immune system, namely Bursa, Peyer's patch, thymus and spleen. Represented on the array are genes that are known to cluster with existing chicken ESTs as well as genes that are unique to our libraries. Some of these genes have no known homologues and represent novel genes in the chicken collection. A series of reference genes (ie. genes of known immune function) are also present on the array. Functional annotation data is also provided for as many of the genes on the array as is possible.

**Conclusion:** Six new chicken immune cDNA libraries have been created and nearly 10,000 sequences submitted to GenBank [GenBank: [AM063043-AM071350](#); [AM071520-AM072286](#); [AM075249-AM075607](#)]. A 5 K immune-related array has been developed from these libraries. Individual clones and arrays are available from the ARK-Genomics resource centre.

### Background

In recent years, the tools available to the field of chicken genomics have increased greatly. Detailed genetic and physical maps have been constructed [1], as well as BAC contig maps [2,3] and a radiation hybrid panel [4]. There is also a substantial EST collection [5], SNP database and

many full-length cDNAs have been sequenced. The development of these resources has culminated with the recent publication of the chicken draft sequence [6]. The chicken can now be regarded as an important model organism for use in comparative genomics, residing in a potentially informative position in the evolutionary ladder. The

chicken is also an extremely useful model for developmental biologists and geneticists as well as being a commercially important species.

The latest tools being developed for the chicken are microarrays. There are several small tissue-specific arrays being used by individual labs. These include an intestinal array (3,072 clones) [7], a macrophage-specific array (4,906 clones) [8], a lymphocyte array (3,011 clones) [9] and an 11 K array based on genes found in heart progenitor cells [10]. A 13 K genome-wide array is also available from ARK-Genomics [11] (Roslin, UK) and from the Fred Hutchinson Cancer Research Centre (Seattle, USA) [12]. We have designed a 5 K immune-related array created from libraries developed from tissue (Bursa, spleen, Peyer's patch, thymus) from birds which were previously inoculated with a combination of different vaccines to various common avian diseases including bacterial, protozoa and virus disease-causing organisms (*E. coli*, Newcastle Disease Virus (NDV), Infectious Bursal Disease Virus (IBDV), coccidiosis, Marek's Disease (MD) and salmonella). The tissues we chose are highly representative of T and B cell populations and were used in order to optimise the numbers of immunologically – related genes that would be present in our libraries. Many known immune genes that have been recently identified in the chicken EST collections [13] have also been added to the array. This array provides a valuable, cost-effective resource for the investigation of immunological gene expression. It has been created from a pool of stimulated immune tissues and contains genes that represent a wide spectrum of immune functions as well as previously unidentified sequences. Each gene on the array is also functionally annotated as much as possible. Gene ontology [14] data and Blast [15] information is provided for each clone, where that information is available.

## Results and discussion

### Construction of the array

Six immune-related libraries were specifically developed for the construction of a 5 K chicken array. Immune tissue from birds inoculated with different vaccine regimes (see Methods) was used to develop two standard libraries. These both underwent two rounds of normalization, thus providing us with six libraries. Initially, 10,173 clones were randomly chosen from the libraries for sequencing. The number chosen from each library depended on the titre (colonies/microlitre) of that particular library. The 10,173 clones that were sequenced were searched for poor quality sequence (<100 bp after removal of vector, repeats etc.) and unwanted Blast homologies, as described in the Methods section. The numbers of high quality sequences (9,434 – which have been submitted to Genbank) from each library are shown in Table 1. Cluster analysis was then undertaken, which resulted in the grouping of clones

**Table 1: Clones sequenced from each library.**

Library	No. of clones
Chicken immune 1 ('B cell' standard)	975
Chicken immune 2 ('B cell' normalized 1)	2,394
Chicken immune 3 ('B cell' normalized 2)	1,600
Chicken immune 4 ('T cell' standard)	2,563
Chicken immune 5 ('T cell' normalized 1)	918
Chicken immune 6 ('T cell' normalized 2)	984

from which we would choose the 5,000 that were to be represented on the array.

### Genes on the array

The clones on the array are derived from custom-made immune-related chicken cDNA libraries. Libraries developed from tissue from Bursa, spleen and Peyer's patch were our representative 'B cell' libraries, and libraries developed from thymus were so-called 'T cell' libraries (the names 'B and T cell libraries' are used purely for ease of reference and in no way indicate that the libraries consist of pure cell populations). Clones from both standard and normalized libraries are present on the array. One clone representing each of the 3,811 clusters is included on the array, along with a random selection of singleton clones (1,067). The sequence of each of the clones was also subjected to a Blast search of the SwissProt and TrEMBL databases and the highest hit to each sequence was reported. Searches were carried out at a stringency of  $1e^{-10}$  (this relatively low stringency was to ensure that we identified as many immune homologies as possible). Chicken immune genes have a relatively low level of sequence conservation with mammals, hence the lower stringency used in these searches). We wanted to ensure that certain genes were also represented on the array as 'reference' genes. This included a range of known immune-related genes for which a clone was already available – either from the existing EST databases [12] or from our novel libraries. Various cytokines, chemokines, cell surface antigens, receptors and MHC molecules were all included (Table 2). The expression profile of genes of unknown function can thus be compared with the profiles of these genes whose roles are known. Standard array controls were also spotted on the array, including various spot report buffers (positive and negative controls for the Cy3 and Cy5 dyes), salmon sperm DNA, calf thymus DNA, bovine genomic DNA (negative controls), chicken genomic DNA, gamma actin and GAPDH (positive controls). Each clone is represented in duplicate.

### Analysis of the immune clones

All the sequences of the clones on the array were subject to Blast homology searches against the SwissProt and TrEMBL databases using a cut-off value of  $1e^{-10}$ . Using

**Table 2: List of known immune genes added as reference genes to the array.**

Gene	EST	Clone ID	Accession no.
<b>AH221</b>	<b>CTN2_C0000858f10.q1kT7SCF</b>	<b>C0000858F10</b>	<a href="#">AM064266</a>
AH294/RANTES	603404971F1	C0000737M17	<a href="#">BU397782</a>
<b>β2 microglobulin</b>	<b>CTst_C0000869a17.q1kT7SCF</b>	<b>C0000869A17</b>	<a href="#">AM068376</a>
<b>BAFF</b>	<b>CBNI_C0000466j11.q1kT7SCF</b>	<b>C0000467J11</b>	<a href="#">AM066201</a>
BMP10	604156553F1	C0003869J14	<a href="#">BU210183</a>
BMP2	603213309F1	C0003763A3	<a href="#">BU444424</a>
BMP4	603363891F1	C0000429F23	<a href="#">BU473912</a>
BMP6	603604307F1	C0003811M1	<a href="#">BU287807</a>
BMP7	603500540F1	C0000717O9	<a href="#">BU333004</a>
BMP8A	603956728F1	603956728F1	<a href="#">BU425800</a>
CC chemokine receptor 6	603508559F1	C0002794E15	<a href="#">BU267610</a>
CC CKR 11	603367511F1	C0000439E3	<a href="#">BU465158</a>
CC LARC/MIP-3A	603534015F1	C0002806N3	<a href="#">BU398190</a>
<b>cCAF</b>	<b>CBNI_C0000465h11.q1kT7SCF</b>	<b>C0000465H11</b>	<a href="#">AM065832</a>
CD135 antigen	603812446F1	C0001334K11	<a href="#">BU376898</a>
CD137	pat.pk0038.d7.f	C0004737E4	<a href="#">AI980851</a>
CD153	pat.pk0072.b5.f	C0004738K22	<a href="#">AI982035</a>
CD154	603535227F1	C0001006F7	<a href="#">BU398104</a>
<b>CD18</b>	<b>CBNI_C0000468e11.q1kT7SCF</b>	<b>C0000468E11</b>	<a href="#">AM066422</a>
CD2	pgn1c.pk014.i9	pk014i9	<a href="#">CB017050</a>
CD200	pat.pk0062.c8.f	C0004738C17	<a href="#">AI981679</a>
CD226	pat.pk0020.d12	C0004739G22	<a href="#">AI980296</a>
<b>CD28</b>	<b>CTst_C0000892d20.q1kT7SCF</b>	<b>C0000892D20</b>	<a href="#">AM070143</a>
<b>CD3</b>	<b>C0001679M3_G02_008.AB1</b>	<b>C0001679M3</b>	<a href="#">AM070515</a>
CD36	603543789F1	C0001028A23	<a href="#">BU311037</a>
CD4	pk017a12	pk017a12	<a href="#">CB017654</a>
CD40L	pgm2n.pk009.d11	pk009d11	<a href="#">BM488880</a>
CD44-like	603745662F1	C0003894K15	<a href="#">BU253134</a>
CD45	603767294F1	C0003827K23	<a href="#">BU446679</a>
CD59	603212850F1	C0000363D13	<a href="#">BU447971</a>
CD63 antigen	603783352F1	C0001268G14	<a href="#">BU243877</a>
CD7	pat.pk0040.d6.f	C0004737K4	<a href="#">AI981043</a>
<b>CD79A</b>	<b>CBst_CHK02000039f07.q1kT7SCF</b>	<b>CHK0200003F7</b>	<a href="#">AM071949</a>
<b>CD8</b>	<b>CTst_C0000877k01.q1kT7SCF</b>	<b>C0000877K1</b>	<a href="#">AM069615</a>
<b>CD82</b>	<b>CTst_C0000892i24.q1kT7SCF</b>	<b>C0000892L24</b>	<a href="#">AM070329</a>
CD83 antigen	603771889F1	C0001238B24	<a href="#">BU457418</a>
<b>CD84</b>	<b>CTNI_C0000798o19.q1kT7SCF</b>	<b>C0000798O19</b>	<a href="#">AM070961</a>
<b>CD98 light chain</b>	<b>CBNI_C0000465c24.q1kT7SCF</b>	<b>C0000465C24</b>	<a href="#">AM065739</a>
Chemokine receptor like 2	603764351F1	C0001219F13	<a href="#">BU444213</a>
Chicken cytokine	pat.pk0050.e9.f	C0004737J11	<a href="#">AI981311</a>
CHIR-A	603478533F1	C0003884A9	<a href="#">BU359209</a>
<b>CHIR-B</b>	<b>CBst_CHK02000039i03.q1kT7SCF</b>	<b>CHK0200003L3</b>	<a href="#">AM072078</a>
cMGF	pat.pk0060.h1.f	C0004737P22	<a href="#">AI981598</a>
<b>Complement C3</b>	<b>CBNI_C0000468j22.q1kT7SCF</b>	<b>C0000468J22</b>	<a href="#">AM066546</a>
Complement C4	603811612F1	C0001332G3	<a href="#">BU376477</a>
Complement C7	603668434F1	C0001140D9	<a href="#">BU416108</a>
Complement C8α	603782386F1	C000164L21	<a href="#">BU242118</a>
complement H	603735023F1	C0001154N6	<a href="#">BU295434</a>
complement receptor 1	603819479F1	C0001351N20	<a href="#">BU268132</a>
<b>Complement I</b>	<b>CBN2_C0000485f23.q1kT7SCF</b>	<b>C0000485F23</b>	<a href="#">AM068133</a>
Cremp	603114782F1	C0003743C21	<a href="#">BU126768</a>
<b>C-type lectin</b>	<b>HFU603551466C18</b>	<b>C0004763C18</b>	<a href="#">AM063354</a>
CX 3C chemokine receptor 1	603949695F1	C0003852N23	<a href="#">BU204148</a>
<b>CXC-R4</b>	<b>CTst_C0000878f17.q1kT7SCF</b>	<b>C0000878F17</b>	<a href="#">AM069849</a>
Cytokine like protein 17	603773283F1	C0001242E21	<a href="#">BU459791</a>
Cytokine receptor like 9	603472805F1	C0000591J1	<a href="#">BU477689</a>
<b>Death receptor 6</b>	<b>CBNI_C0000466i11.q1kT7SCF</b>	<b>C0000467L11</b>	<a href="#">AM066244</a>
DSL-1	603321647F1	C0000418M11	<a href="#">BU239031</a>
EMAP 11	603364164F1	C0003776P16	<a href="#">BU475067</a>

**Table 2: List of known immune genes added as reference genes to the array.** (Continued)

ephrin type A receptor 2	603121949F1	C0000241A13	<a href="#">BU133519</a>
FAS antigen	603737578F1	C0001181L19	<a href="#">BU300974</a>
<b>FASL decoy receptor 3</b>	<b>CTN2_C0000856k13.q1kT7SCF</b>	<b>C0000856K13</b>	<a href="#">AM064070</a>
<b>GATA-3</b>	<b>CTN2_C0000858a24.q1kT7SCF</b>	<b>C0000858A24</b>	<a href="#">AM064179</a>
GDF10	603530236F1	C0000994G10	<a href="#">BU351257</a>
GDF8	603775823F1	C0001248N6	<a href="#">BU458566</a>
GDF9	603741256F1	C0001166O13	<a href="#">BU300398</a>
glycoprotein I30	603369157F1	C0002739A15	<a href="#">BU460413</a>
GMCSF	CF258055	CF258055	<a href="#">CF258055</a>
HCC1	pat.pk0059.g4.f	pk059g4	<a href="#">AVV061438</a>
ICSBP	603568552F1	C0001037G14	<a href="#">BU383423</a>
IFN $\alpha$	603486811F1	C0000622F3	<a href="#">BU319434</a>
IFN $\alpha$ / $\beta$ -R2	603783234F1	C0001268C9	<a href="#">BU243612</a>
IFN $\gamma$	603766180F1	C0003825O20	<a href="#">BU444142</a>
IFN $\gamma$ R2	603606133F1	C0001120B10	<a href="#">BU294744</a>
IFP35	603123028F1	C0000243H7	<a href="#">BU135331</a>
<b>Ig light chain VJC region</b>	<b>C0000914E7_C03_018.AB1</b>	<b>C0000914E7</b>	<a href="#">AM064528</a>
Ig heavy chain VDJ region	603534767F1	C0002807F18	<a href="#">BU398082</a>
IK cytokine	603368212F1	C0000440H3	<a href="#">BU460192</a>
IL-10	CF258071	CF258071	<a href="#">CF258071</a>
IL11 receptor	603402722F1	C0000518K5	<a href="#">BU250398</a>
IL-12 $\beta$	603603708F1	603603708F1	<a href="#">BU291084</a>
IL12-p35	603761859F1	C0002846F16	<a href="#">BU474924</a>
IL-13R2	603519773F1	C0000972A19	<a href="#">BU341330</a>
IL-15	603102514F1	C0002655L4	<a href="#">BU202444</a>
IL-16	603130176F1	C0002702H20	<a href="#">BU114872</a>
IL17 receptor	603211483F1	C0000350E22	<a href="#">BU448712</a>
IL-18	603508766F1	C0002794D18	<a href="#">BU271029</a>
IL-1 $\beta$	603217760F1	C0003766G15	<a href="#">BU455380</a>
IL-2	pat.pk022.e2	pk022e2	<a href="#">AI980311</a>
IL20 receptor	603591538F1	C0001088K9	<a href="#">BU241765</a>
IL-2R $\alpha$	pat.pk0012.h3	pk0012h3	<a href="#">AI980106</a>
<b>IL-2R<math>\gamma</math></b>	<b>CBNI_C0000360j15.q1kT7SCF</b>	<b>C0000360J15</b>	<a href="#">AM064841</a>
IL-4	603772775F1	ChEST708F19	<a href="#">BU450270</a>
IL-4R	603490820F1	603490820F1	<a href="#">BU324362</a>
IL-6	pat.pk0076.f2.f	C0004739G21	<a href="#">AI982185</a>
Interleukin enhancer binding factor 3	603322645F1	C0000420J12	<a href="#">BU239448</a>
IRAK2	603831145F1	ChEST821j11	<a href="#">BU435261</a>
IRAK4	603208981F1	ChEST185a21	<a href="#">BU441365</a>
IRF1	603960463F1	C0002900N15	<a href="#">BU418343</a>
<b>IRF10</b>	<b>CBNI_C0000360i15.q1kT7SCF</b>	<b>C0000360L15</b>	<a href="#">AM064884</a>
IRF2	604146465F1	C0003862A18	<a href="#">BU438609</a>
<b>IRF3</b>	<b>CTst_C0000892j09.q1kT7SCF</b>	<b>C0000892J9</b>	<a href="#">AM070266</a>
<b>IRF4</b>	<b>604_B10_077</b>	<b>C0000885O4</b>	<a href="#">AM072251</a>
IRF5	pat.pk0067.c5.f	pk067c5	<a href="#">AI981854</a>
IRF6	603111427F1	C0000188F7	<a href="#">BU109331</a>
<b>JSC</b>	<b>CTst_C0000878m23.q1kT7SCF</b>	<b>C0000878M23</b>	<a href="#">AM069996</a>
K60	603470605F1	C0002774I2	<a href="#">BU479398</a>
lymphotactin	603733847F1	C0001151N12	<a href="#">BU300469</a>
MCSF1-receptor	603220574F1	C0000383C19	<a href="#">BU432910</a>
<b>MDV vIL-8</b>	<b>CBst_C0000222p17.q1kT7SCF</b>	<b>C0000222P17</b>	<a href="#">AM071831</a>
<b>MHC class I</b>	<b>CBst_CHK02000039o05.q1kT7SCF</b>	<b>CHK0200003O5</b>	<a href="#">AM072147</a>
<b>MHC class I minor</b>	<b>CTst_C0000873a15.q1kT7SCF</b>	<b>C0000873A15</b>	<a href="#">AM068728</a>
<b>MHC class II beta</b>	<b>HFU603551341A11</b>	<b>C0004763A11</b>	<a href="#">AM063247</a>
MIF	604141521F1	C0001517I5	<a href="#">BU438017</a>
MX	603775783F1	C0001248E22	<a href="#">BU457953</a>
NKL	603539011F1	C0001016E4	<a href="#">BU309556</a>
N-pac	604157079F1	604157079F1	<a href="#">BU210048</a>
NRAMP1	pk013p5	pk013p5	<a href="#">BI394251</a>
NRAMP2	603953027F1	C0001451E8	<a href="#">BU203948</a>
OCIF	603157972F1	C0000333O17	<a href="#">BU410189</a>
opioid receptor sigma 1	603341826F1	C0003775C8	<a href="#">BU254440</a>
Orphan chemokine receptor	603234142F1	C0000403C12	<a href="#">BU418544</a>

**Table 2: List of known immune genes added as reference genes to the array.** (Continued)

PIT54	603150061F1	C0000313116	<a href="#">BU126277</a>
platelet activating receptor	pat.pk0002.b12	C0004739A2	<a href="#">AI979750</a>
PRC1	603475588F1	603475588F1	<a href="#">BU355757</a>
<b>prostaglandin synthase</b>	<b>CBNI_C0000466j02.q1kT7SCF</b>	<b>C0000467J2</b>	<a href="#">AM066193</a>
Regulator of cytokinesis I	603475588F1	C0000598P22	<a href="#">BU355757</a>
<b>RING3</b>	<b>CBst_C0000222p04.q1kT7SCF</b>	<b>C0000222P4</b>	<a href="#">AM071819</a>
<b>SCA-2</b>	<b>CTNI_C0000853f13.q1kT7SCF</b>	<b>C0000853F13</b>	<a href="#">AM071100</a>
SCYa13	603534566F1	603534566F1	<a href="#">BU397023</a>
SCYA4	603742061F1	C0001168115	<a href="#">BU299262</a>
SIGIRR	603321436F1	C0002934H10	<a href="#">BU240159</a>
SOCS1	603758706F1	C0003823K20/L1	<a href="#">BU218362</a>
SOCS2	603322984F1	603322984F1	<a href="#">BU239208</a>
SOCS5	603492126F1	C0000636C7	<a href="#">BU326390</a>
STAT1	603957345F1	ChEST927111	<a href="#">BU425993</a>
STAT2	pat.pk0027.a6.f	C0004737C3	<a href="#">AI980571</a>
<b>STAT5b</b>	<b>CBst_C0000222j02.q1kT7SCF</b>	<b>C0000222J2</b>	<a href="#">AM071688</a>
TAP2	603732809F1	C0000758A3	<a href="#">BU298074</a>
<b>Tapasin</b>	<b>CTst_C0000873d22.q1kT7SCF</b>	<b>C0000873D22</b>	<a href="#">AM068799</a>
TARC	pat.pk0031.f10.f	C0004737K13	<a href="#">AI980713</a>
T-bet	pgn1c.pk013.h8	pgn1c.pk013.h8	<a href="#">CB016768</a>
<b>Tcell receptor <math>\alpha</math></b>	<b>CBst_CHK02000039p10.q1kT7SCF</b>	<b>CHK0200003P10</b>	<a href="#">AM072172</a>
<b>Tcell receptor <math>\beta</math></b>	<b>UEB603581072O18</b>	<b>C0004765O18</b>	<a href="#">AM063532</a>
<b>Tcell receptor <math>\gamma</math></b>	<b>CTst_C0000878m08.q1kT7SCF</b>	<b>C0000878M8</b>	<a href="#">AM069981</a>
<b>Tcell receptor <math>\zeta</math></b>	<b>CTst_C0000874j17.q1kT7SCF</b>	<b>C0000874J17</b>	<a href="#">AM069278</a>
TGF $\beta$	603758578F1	C0003823K20	<a href="#">BU215243</a>
<b>Thymosin beta 4</b>	<b>ODP603945810C04</b>	<b>C0004766C4</b>	<a href="#">AM063804</a>
TLR1/6/10	603760940F1	C0001211O10	<a href="#">BU471724</a>
TLR2	603588755F1	C0001081M13	<a href="#">BU374739</a>
TLR3	603781018F1	C0001261D6	<a href="#">BU242827</a>
TLR4	603470778F1	C0002774L20	<a href="#">BU475859</a>
TLR5	603230983F1	C0000395E22	<a href="#">BU420247</a>
TLR7/8/9	603160284F2	C0002711G22	<a href="#">BU435893</a>
TRAF1	pat.pk0072.d3.f	C0004738M6	<a href="#">AI982046</a>
TRAF2	603217872F1	C0003766O18	<a href="#">BU455745</a>
<b>TRAF5</b>	<b>CTst_C0000877n08.q1kT7SCF</b>	<b>C0000877N8</b>	<a href="#">AM069687</a>

The genes in bold come from the immune libraries described in this paper

this means of detection, many known immune-related molecules were identified, including cytokines, interferons, interleukins, transcription factors, receptors, cell differentiation antigens, MHC molecules and genes for proteins belonging to the TOLL receptor pathway. Proteins homologous to hypothetical human proteins and mouse cDNAs were also identified.

Sequences, which gave no Blast homology to anything in the nucleotide or protein databases, accounted for about 38% of the clones. Either the search parameters were too stringent to identify these genes or the chicken sequence was sufficiently divergent to be undetectable in a standard Blast search. This is a common feature of immune-related genes, and it is often very difficult to identify such genes by sequence homology to mammalian homologues. Some of these sequences may also represent non-conserved 3' UTR regions of genes. This set of clones may also include genes that have never been identified before and are not represented in the sequence databases. Further,

more detailed analysis of these sequences can sometimes help elucidate the nature of the gene in question. Protein sequences can be predicted from the EST nucleotide sequence using programs such as ESTscan [[16] and [17]], which takes in to account sequencing errors and thus potential frame-shift mutations which are often present when there is only one EST sequence available for study. Conserved motifs and domains can then sometimes be identified for example, using the Pfam database [18], which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains and families. PSI-Blast searches can also help identify to which type of family a gene will belong.

During clustering analysis, our 10,000 immune sequences were aligned with 398,000 existing chicken ESTs. This highlighted 3,845 clusters that contained one or more sequence from our immune libraries and 1,959 singleton clones. This analysis also identified 40 novel clusters that

**Table 3: Genomic location of unique chicken ESTs as identified by the University of Santa Cruz Blat site <http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>**

Clone	GenBank Accession No.	Genomic location
CBN1_C0000465p01.q1kT7	<a href="#">AM065989</a>	chr15: 5562143–5562781
CBN1_C0000464a07.q1kT7	<a href="#">AM065333</a>	chr11_random: 637442–638117
CBN1_C0000465g01.q1kT7	<a href="#">AM065802</a>	chr28: 3396525–3399444
CBN2_C0000485f09.q1kT7SCF	<a href="#">AM068122</a>	chr7: 1319368–1319989
CBst_C0000222i16.q1kT7SCF	<a href="#">AM071684</a>	no hit
CTN2_C0000856o09.q1kT7	<a href="#">AM064132</a>	chr17: 9342673–9343272
CBN1_C0000360j04.q1kT7	<a href="#">AM064831</a>	chr21: 2494538–2494813
CBN1_C0000360n18.q1kT7	<a href="#">AM064932</a>	chr1: 172893324–172893990
CBN1_C0000463a02.q1kT7	<a href="#">AM064982</a>	chr4: 50741781–50742197

only contained sequences from our new libraries. Upon Blast analysis, 7 of these clusters were found to represent known chicken genes (initially appearing unique as they aligned to a different part of the gene sequence from existing ESTs), 18 showed homology to genes in other species and 15 clusters proved to have no known homology to anything currently in the databases. At the time, we searched against 398,000 existing chicken ESTs. Now however, there are currently 550,510 chicken ESTs in the databases (dbEST release 080505). A current search has shown that 9 of our sequences are indeed still unique to our libraries and have no known identifiable homologue, although two of the sequences do show some similarity to two predicted chicken sequences ([AM065333](#) and the hypothetical protein [XP\\_429359](#); [AM065802](#) and the predicted P114-RHO-GEF protein [XP\\_418249](#)). Eight of these sequences are identifiable in the whole genome sequence, as shown in Table 3.

#### Gene ontology (GO) annotations

In order to try and elucidate the function of the genes on the array further, we tried to assign as much annotation to the sequences as possible. GO annotations were assigned to some sequences after searching the GGI and UMIST databases [19], while other annotation was derived from hits to orthologous human sequences from the ENSEMBL [20] and GENSCAN [21] databases, as described in the 'methods' section. Having annotation derived from orthologous human genes means that cross-species comparisons between chicken and human array data may be possible. A search of the ENSEMBL database provided information on 2,292 GO-term associations, the GGI database 1,542 and GENSCAN 566, while the UMIST full-length cDNA database provided a further 365 annotations. The sequences on the array cover a total of 227 GO terms, with 73% of all the sequences having at least one GO entry assigned to it. The available annotation for the array sequences is broken down as follows: 52% of genes have a 'cellular component' term assigned, 60% have 'molecular function' and 56% of sequences have the 'biological process' described. 83% of all the genes on the

array have some kind of gene description and after searching each sequence against the sequences in the Ensembl chicken genome collection (July 2005 genebuild [22]), 78% of sequences were found to have a known chromosomal location. Now that all these sequences have been added to GenBank and thus have an accession number which can be directly linked into the ENSEMBL databases (work currently underway), obtaining comprehensive, up-to-date annotation data will become much easier.

A file showing the complete annotation for all the sequences on the array is available as supplementary material (Additional file 1). However, Additional file 2 provides an overview of the broad functional classes that are represented by the genes on the array. These are based on more general GO annotations derived from the GOslims database at EBI, and allow us an insight into the different classes of genes present on the array without having to look at detailed functional annotation for each individual gene.

Annotation is also available for some (9,137) of the ESTs in the UMIST collection. By comparing the relevant GOslims [23] terms for the sequences in this collection with those present on our array, we are able to see which types of genes appear to be enriched in our set, compared with a larger, more general collection of EST sequences. As can be seen (shown in bold) in Additional file 2, certain classes of gene appear to be more highly represented. For instance, genes involved in protein transport are more abundant in our set of clones, as are those involved in the response to stimulus. This is consistent with our attempts to pre-select for higher numbers of genes involved in the immune system.

#### Quality of the array

To assess the quality of the array, various hybridization comparisons were undertaken. Three different conditions were addressed: 1). self v self 2). biological replicate A v biological replicate B and 3). Control sample v activated sample. Dye swap experiments were also carried out for

conditions 2 and 3. The 'self' sample was a reference RNA consisting of a pool of various chicken lung samples. The biological replicates were lung samples from two 6-week-old chickens that had not been treated or challenged in any way. In the third group of hybridisations, the 'control' sample was from a similarly, untreated bird and the 'activated' sample was obtained from the lungs of a bird that had been challenged with the avian influenza strain H9N2 five days previously. The graphs in Fig 1 show the tight correlation between self/self ( $R^2 = 0.9273$ ) and between replicates ( $R^2 = 0.8766$ ), whereas a much higher level of variance is seen when an activated sample is compared against a control ( $R^2 = 0.7601$ ).

The boxplots in Fig 2 also demonstrate the differing variances between the comparisons. The greatest variance is shown for the activated animals compared with the controls as would be expected. Regression analysis for each of the data sets confirm the increased variance with correlation coefficients of  $r = 0.872$  for activated samples,  $r = 0.936$  for replicate samples and  $r = 0.963$  for self/self sample data sets.

#### Using the array

This array is available from the Ark-Genomics resource facility at Roslin Institute, providing an immune-focused array which, for anyone interested in immune-research, offers a much more cost-effective and time-saving platform for gene expression experiments, instead of using the large oligo arrays which have thousands more genes, many of which will be of no interest. Analysis of data is also thus much easier and far less time-consuming. Information on the array has been deposited in ArrayExpress (Accession: A-MEXP-307) [[24] and [25]] (Additional file 1) and very soon all the sequences will be submitted to the Ensembl database with links to all the GO annotation information in the GOA database [26].

#### Conclusion

We have constructed a 5 K chicken cDNA microarray, which is highly selected for genes expressed in tissues which have an immune function. This targeted array contains enough widely-expressed genes (whose expression won't be changing) to enable good normalization, as well as containing numerous known immune genes (from our novel libraries and from existing EST collections). The array also contains many genes with as yet unknown homology and function as well as a few novel genes which are specific to the libraries from which the array was created. These genes of unknown function could well have a role in either the adaptive or innate immune response, and thus provide a valuable resource for analysis of gene expression changes occurring in birds that have been subject to immune challenge. The array has been proven to

provide highly reproducible results and is now available to the chicken/microarray community as a whole.

#### Methods

##### Sample collection

Eight groups of 38 chickens (3-week-old) were vaccinated with two different vaccine regimes. The eight groups were males and females of a commercial line of hybrid broiler (Ross 306, Aviagen, Newbridge, Midlothian, UK) and layer (Lohman Brown, Lohmann Tierzucht, Cuxhaven Germany) chicks given one of the two vaccination schemes. Group 1 were given vaccines for *E. coli* (0.5 ml in left breast muscle), ND and IBDV (0.5 ml in right breast muscle) formulated in alum-gel and oil-based immunopotentiators. Intramuscular injections were given to ensure that all the birds were given an equal dose. Group 2 vaccines consisted of Paracox 8 [*Eimeria* sp.] (0.1 ml in drinking water), Nobilis Rismavac-CA126 [MD] (0.2 ml intramuscularly in leg) and Salenovac [*S. enteritidis*] (0.5 ml intramuscularly in leg). Tissue samples were obtained (unvaccinated); 5 hr, 24 hr, 72 hr and 7 days post vaccination. Samples from groups of 5 birds were pooled. Tissues collected were Bursa, spleen, Peyer's patch and thymus. Tissue from Bursa, spleen and Peyer's patch were pooled to make the 'B-cell' libraries and the thymus tissue was used to construct the 'T-cell' libraries. The tissues and time points chosen were in order to try and maximise the number of immune-related transcripts, including those which may only be expressed transiently. All experimental protocols were authorized under the UK Animals (Scientific Procedures) Act, 1986.

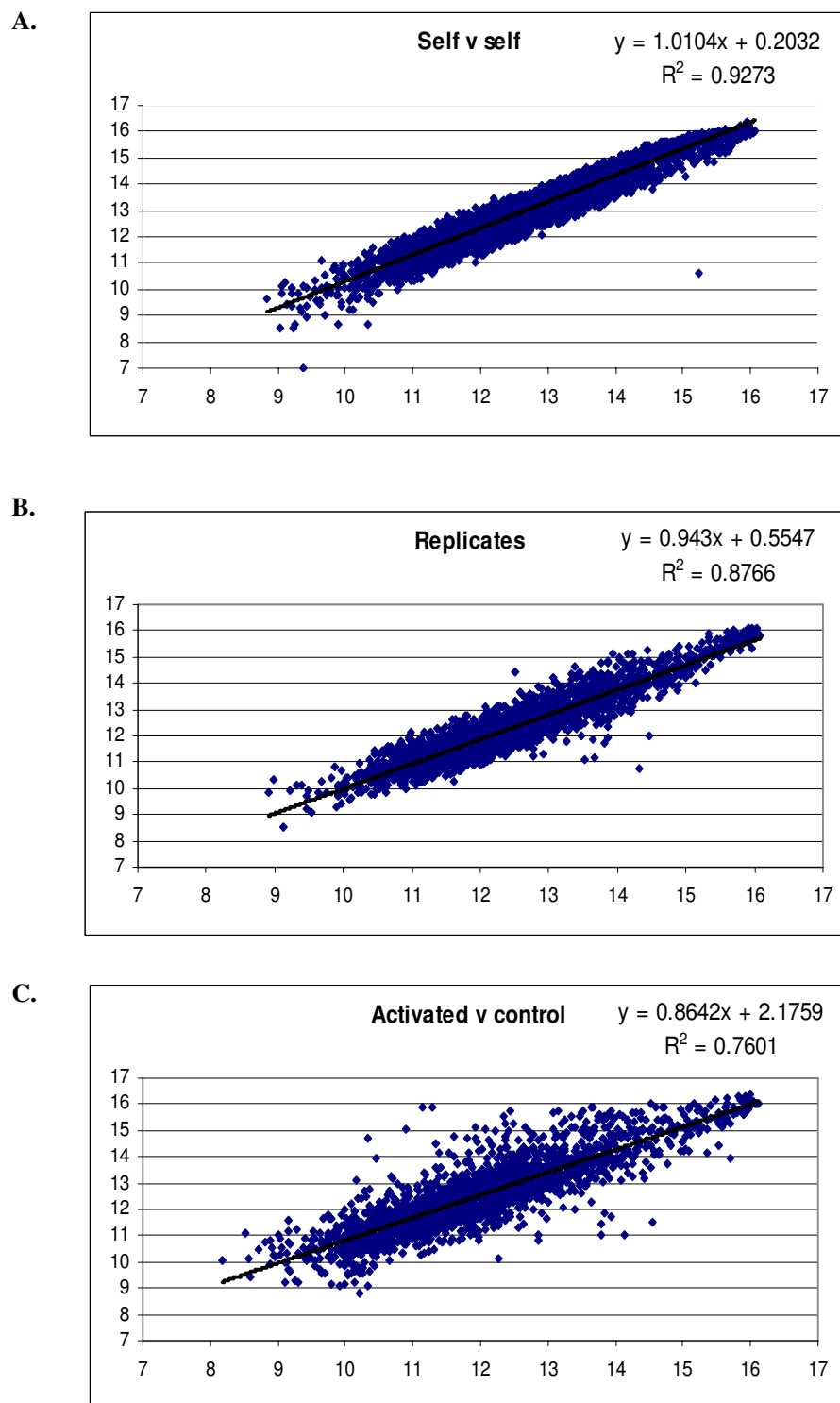
##### Library construction

Six libraries were constructed at Incyte Genomics (Palo Alto, CA): a standard and 2 normalized Bursa/spleen/Peyer's patch libraries and a standard and 2 normalized thymus libraries. cDNA synthesis was initiated using an oligo (dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to *NotI* adapters, digested with *EcoRI*, size-selected, and cloned into the *NotI* and *EcoRI* compatible sites of a custom modified MCS of the pBluescript (KS+) vector. Normalization was done in two rounds using conditions adapted from [[27] and [28]] except that a significantly longer re-annealing hybridization was used. Around 10,000 clones were then sequenced at the Sanger Institute according to their protocols. Using the T7 primer, sequence was generated from the 5' end of each clone by the dideoxy chain termination method using an ABI 3700 sequence analyser (Applied Biosystems, Foster City, CA).

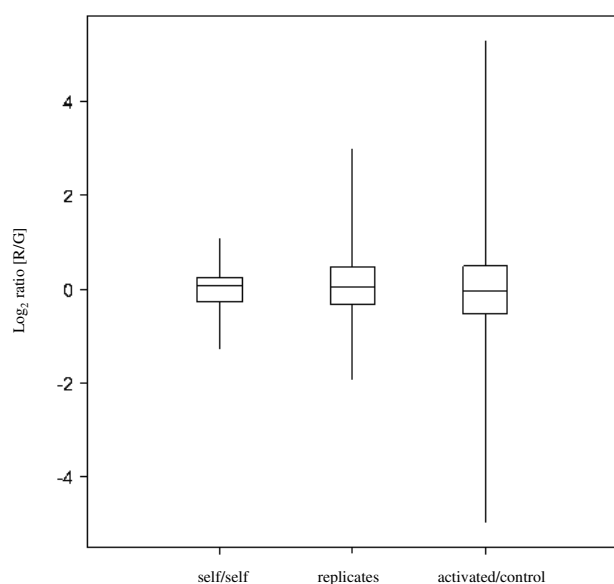
##### EST sequence analysis

Bioinformatic analysis commenced with 10,173 sequences. After eliminating poor quality sequence and



**Figure 1**

Scatter plots showing the variance between A). self/self hybridisation B). two biological replicates and C). a control sample compared with an activated sample. Very little spread is seen with the self/self hybridisation and between the two replicates, as would be expected. However, differences in gene expression can be seen between the activated and control samples.



**Figure 2**  
Box plots showing the variance between self/self hybridisation, two biological replicates and a control sample compared with an activated sample. Boxes represent the interquartile range from 25–75%, with the median marked. Outliers to this range are also shown.

repeats, 9,434 of these sequences remained after screening with phred [29], RepeatMasker [30], Crossmatch [31] and XNUN [32]. Certain unwanted sequences were then identified after using the Blast algorithm [[33] and [34]] and screening the results for specific keywords. These included 'ribosomal', 'mitochondrial', 'Newcastle', 'Mareks', 'Eimeria', 'Salmonella' and 'E. coli'. 8,154 sequences passed these criteria. These sequences were then clustered against the existing UMIST and EMBL chicken EST sequences using TIGR's clustering tool, tgc1 [35]. This resulted in 3,845 clusters which contained one or more sequence from our libraries and 1,959 singletons. The following clones were chosen for inclusion on the array: 3,770 cluster representatives, 1,067 singletons and 157 reference immune genes: 93 clones from the UMIST collection, 41 from our immune libraries, 21 clones from the Delaware set [36] and 2 clones courtesy of R. Zoorob (CNRS, France) (Table 2).

#### Construction of the array

The immune array was constructed from 4994 chicken EST clones plus 196 control elements (landing lights (positional controls), GAPDH, gamma actin, salmon sperm DNA, calf thymus DNA, chicken and bovine genomic DNA and a variety of spotting buffers). Plasmid

DNA was prepared using MagAttract 96 Miniprep chemistry on a Biorobot 8000 platform (Qiagen Ltd., Crawley, UK), and the cDNA inserts were amplified using CGAT-TAAGTTGGGTAACGC (fwd) and CAATTCACACAG-GAAACAG (rev) in 50  $\mu$ l reactions using 1  $\mu$ l of DNA as a template. Amplified DNA was purified by Multiscreen 384 well PCR purification plates (Millipore, Watford, UK) on a Multiprobe II liquid handling platform (Perkin Elmer, Beaconsfield, UK) and the reactions confirmed by agarose gel electrophoresis and quantified by Picogreen assay (Molecular Probes, Invitrogen, Paisley, UK) on a Flouroskan Ascent fluorescent plate reader (Thermo Life Science, Basingstoke, UK). DNA was resuspended to 150 ng/ $\mu$ l in spot buffer (150 mM Sodium phosphate, 0.01% SDS) before being spotted in duplicate on to amino-silane coated slides (CMT-GAPSII, Corning, Schiphol-Rijk, The Netherlands) using a Biorobotics MicroGrid II spotter (Genomic Solutions, Huntingdon, UK). Slides were then treated using succinic anhydride and 1-methyl-2-pyrrolidinone (Sigma, Poole, UK) to block unbound amino groups, followed by a wash in 95°C MilliQ water before hybridisation.

#### RNA preparation and labelling

Total RNA was isolated from lung tissue using a Trizol extraction according to the manufacturer's protocol (Invitrogen, Paisley, UK) and subsequently purified using the RNeasy Midi RNA Purification kit (Qiagen Ltd., Crawley, UK). RNA concentration was determined spectrophotometrically and RNA quality was determined using an Agilent 2100 Bioanalyser (Agilent Technologies, Waldbronn, Germany). Cy3 or Cy5 was incorporated into each sample using the Fairplay labelling kit (Stratagene, La Jolla, CA) and the labelled cDNA cleaned-up after passage through DyeEx columns (Qiagen Ltd., Crawley, UK). Labelling efficiency was determined by running 0.5  $\mu$ l of each sample on a 1% agarose gel and measuring the intensity of fluorescence on a GeneTac LS IV scanner (Genomic Solutions, Huntingdon, UK).

#### Hybridizations

Microarray hybridizations were carried out overnight using a GeneTAC automated hybridization system [37] (Genomic Solutions, Huntingdon, UK). Hybridizations (125  $\mu$ l) were carried out in Genomic Solutions hybridization solution (Cat. no. RP#0025) in a stepped hybridization: 55°C for 3 hr, 50°C for 3 hr and then 45°C for 12 hr. Slides were then washed in Genomic Solutions wash buffers (Cat. nos. CS#0038, CS#0039 and CS#0040). Upon removal from the hybridization stations, slides were washed for 1 min in Post-Wash buffer (CS#0040) and a further minute in isopropanol, followed by centrifugation at 1000 rpm for 6 min. Dried slides were scanned in a Scanarray 5000 scanner (GSI Lumonics, Rugby, UK) fitted with Cy3 and Cy5 filters.

### Data analysis

To indicate the suitability of the new array to discriminate the differences in the experimental treatments, hybridizations comparing samples with controls and controls with controls were performed. Control (vehicle treated) animals were compared with immunologically challenged animals (activated slides) and control animals were also compared with other control individuals (replicate slides). The same animal was also compared with itself (self/self). Each comparison was completed in duplicate and with a dye flip. Dye-swaps are carried out in order to deal with any residual dye-bias remaining after labelling. However, this is generally not a problem, due to the indirect labelling method employed. Data was extracted from the slide using Bluefuse software (BlueGnome, Cambridge, UK). Features with poor confidence information (confidence < 0.30, flagged D and E) were eliminated from the analysis.  $M \text{ v } A$  plots [where  $M = \log_2(\text{Cy5/Cy3})$  and  $A = 1/2 * (\log_2(\text{Cy5}) + \log_2(\text{Cy3}))$ ] of the data for each slide (data not shown) were suitably linear to require only a simple global normalisation of the data. Data from slides of similar treatments was pooled and a boxplot produced for each comparison (Genstat v8.1, VSN International Ltd., Hemel Hempstead, Herts, UK).

### Databases and sequence sources

Ensembl and Genscan predicted genes/peptide sequences for the chicken genome assembly (March 2004) were downloaded from the Ensembl database using Ensmart or the UCSC table browser [38]. Chicken EST sequences were downloaded from the TIGR *Gallus gallus* gene index (GGGI) [release 10.0] [39] and [40]. Chicken full-length cDNA sequences were downloaded from the UMIST www site (Sept 2004). Ensembl predicted peptide sequences for the human genome assembly (May 2004) were downloaded from the Ensembl database using Ensmart or the UCSC table browser.

### Mapping array probes to chicken ESTs, cDNAs, genes and genome

Unique ESTs used to create the immune array were mapped to chicken cDNAs, ESTs, genes or the chicken genome assembly using NCBI Blastn (version 2.2.11). Identity was defined with > 95% sequence identity over 100-bp and then taking the top-scoring match to each EST to provide a unique sequence assignment. All repeats and low-complexity sequences were masked using RepeatMasker (version 3.1.0).

### Definition of Gene Ontology terms and Gene Descriptions for array probes

Gene Ontology (GO) annotations [41] were all based on database hits in sequence similarity searches using Blastn. GO annotations were automatically transferred from these database records to the array probe entries. GO

annotations were available for GGGI and UMIST EST/cDNA sequences. For chicken Ensembl or Genscan gene predictions, GO annotations were based on orthologous human peptide sequences. Orthologues were defined based on two cycles of Blastp between human and chicken proteins. An E\_value cut off of less than  $10^{-4}$ , with the subject and query databases swapped between runs. By comparing E\_values mutually best proteins pairs were selected as orthologues. When E\_values were equal, bits score and sequence coverage were used as tiebreakers to select the top hit. For each array probe associated GO terms and a unique gene description was transferred from the orthologous database record. Finally a Perl script was used to create a non-redundant set of probe to GO records.

### Frequency of GO and GO-Slim terms

GO terms (version 3.2.16) were downloaded from the Gene Ontology www site. More general GO terms were assigned using GoSlim\_map (June 2005) available from the GOA www site at EBI. The GO-Slim terms allowed us to estimate e.g. the frequency of array probes associated with the biological process *Metabolism* (GO:0008152).

### Data processing

Perl scripts (version 5.8.5) and SQL were used throughout to manipulate and filter data sets.

### Authors' contributions

JS contributed to the design of the array, carried out the microarray experiments and drafted the manuscript. DS carried out quality control, clustering and BLAST searches of the DNA sequences. PH was responsible for tissue sample collection. RTT statistically analysed the microarray data. WGJD and VEJCS were involved in experimental design. EJG contributed to the array design. DWB carried out all the bioinformatics involved in establishing GO annotations. All authors read and approved the final manuscript.

### Additional material

#### Additional File 1

The file *supp\_mat.xls* is an excel file which contains annotation information on the sequences present on the immune array and is available as supplementary material. An ArrayExpress <http://www.ebi.ac.uk/arrayexpress> file is available under accession number A-MEXP-307.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-49-S1.xls>]

## Additional File 2

The file *supp\_mat\_2.xls* is an excel file and contains a summary of functional groups present on the array (GO slims). Percentages are calculated as a fraction of the total number of classes represented within one functional description. For example, 42.4% of all the genes are involved in some kind of physiological process. Of these, 14.3% are involved in transport, with 10.2% of these genes being specifically involved in electron transport. This breakdown of functional classes is compared to those represented by 9,137 of the ESTs in the UMIST collection (data available from <http://chick.umist.ac.uk>). Entries shown in bold define the GO classifications that appear to be enriched in the sequences represented on the immune array compared with this subset of the UMIST chicken ESTs. Click here for file  
[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-49-S2.xls>]

## Acknowledgements

The authors would like to thank Incyte Genomics (Palo Alto, CA) for construction of the normalized cDNA libraries and The Wellcome Trust Sanger Institute (Hinxton, UK) for sequencing 10,000 cDNA clones from the libraries. Thanks also to Frazer Murray of ARK-Genomics (Roslin) for invaluable technical assistance and to Theo Jansen (Intervet International B.V., Boxmeer, The Netherlands) for the preparation of the vaccine formulations. This project was funded by Intervet International B.V., Boxmeer, The Netherlands, the Biotechnology and Biological Science Research Council (BBSRC) and partly by a BSIK VIRGO consortium grant, the Netherlands (grant nr. 03012).

## References

- Schmid M, Nanda I, Guttenbach M, Steinlein C, Hoehn M, Scharl M, Haaf T, Weigend S, Fries R, Buerstedde JM, Wimmers K, Burt DW, Smith J, A'Hara S, Law A, Griffin DK, Bumstead N, Kaufman J, Thomson PA, Burke T, Groenen MA, Crooijmans RP, Vignal A, Fillon V, Morisson M, Pitel F, Tixier-Boichard M, Ladjali-Mohammedi K, Hillel J, Maki-Tanila A, Cheng HH, Delany ME, Burnside J, Mizuno S: **First report on chicken genes and chromosomes 2000**. *Cytogenet Cell Genet* 2000, **90**:169-218.
- Aerts J, Crooijmans R, Cornelissen S, Hemmatian K, Veenendaal T, Jaadar A, van der Poel J, Fillon V, Vignal A, Groenen M: **Integration of chicken genomic resources to enable whole-genome sequencing**. *Cytogenet Genome Res* 2003, **102**:297-303.
- Ren C, Lee MK, Yan B, Ding K, Cox B, Romanov MN, Price JA, Dodgson JB, Zhang HB: **A BAC-based physical map of the chicken genome**. *Genome Res* 2003, **13**:2754-2758.
- Morisson M, Lemiére A, Bosc S, Galan M, Plisson-Petit F, Pinton P, Delcros C, Feve K, Pitel F, Fillon V, Yerle M, Vignal A: **ChickRH6: a chicken whole-genome radiation hybrid panel**. *Genet Sel Evol* 2002, **34**:521-533.
- Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E, Fong WT, Tickle C, Brown WR, Wilson SA, Hubbard SJ: **A comprehensive collection of chicken cDNAs**. *Curr Biol* 2002, **12**:1965-1969.
- International Chicken Genome Sequencing Consortium (ICGSC): **Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution**. *Nature* 2004, **432**:695-716.
- van Hemert S, Ebbelaar BH, Smits MA, Rebel JM: **Generation of EST and microarray resources for functional genomic studies on chicken intestinal health**. *Anim Biotechnol* 2003, **13**:133-143.
- Bliss TW, Dohms JE, Emara MG, Keeler CL Jr: **Gene expression profiling of avian macrophage activation**. *Vet Immunol Immunopath* 2005, **105**:289-299.
- Neiman PE, Ruddell A, Jasoni C, Loring G, Thomas SJ, Brandvold KA, Lee RM, Burnside J, Delrow J: **Analysis of gene expression during myc oncogene-induced lymphomagenesis in the bursa of Fabricius**. *Proc Natl Acad Sci (USA)* 2001, **98**:6378-6383.
- Afrakhte M, Schultheiss TM: **Construction and analysis of a subtracted library and microarray of cDNAs expressed specifically in chicken heart progenitor cells**. *Dev Dynam* 2004, **230**:290-298.
- ARKGenomics [<http://www.ark-genomics.org>]
- Burnside J, Neiman P, Tang J, Bascom R, Aronszajn M, Talbot R, Burt DW, Delrow J: **Development of a cDNA array for chicken gene expression analysis**. *BMC Genomics* 2005, **6**:13.
- Smith J, Speed D, Law AS, Glass EJ, Burt DW: **In silico identification of chicken immune-related genes**. *Immunogenetics* 2004, **56**:122-133.
- Gene ontology www site [<http://www.geneontology.org/>]
- Blast at NCBI [<http://www.ncbi.nlm.nih.gov/BLAST/>]
- Iseli C, Jongeneel CV, Bucher P: **ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences**. *Proc Int Conf Intell Syst Mol Biol* 1999:138-148.
- ESTscan [<http://www.ch.embnet.org/software/ESTScan.html>]
- Pfam [<http://www.sanger.ac.uk/Software/Pfam/>]
- BBSRC chicken EST database [<http://chick.umist.ac.uk/>]
- Ensembl genome databases [<http://www.ensembl.org/>]
- Genscan [<http://genes.mit.edu/GENSCAN.html>]
- Ensemble 2005 chicken genebuild [<ftp://ftp.ensembl.org/pub/chicken-32.1h/data/fasta/dna/>]
- GO slims [<http://www.geneontology.org/GO.slims.shtml>]
- Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, Holloway E, Kapushesky M, Kemmeren P, Lara GG, Oezcimen A, Rocca-Serra P, Sansone S: **ArrayExpress - a public repository for microarray gene expression data at the EBI**. *Nucleic Acids Res* 2003, **31**:68-71.
- ArrayExpress [<http://www.ebi.ac.uk/arrayexpress/>]
- Gene ontology annotation at EBI [<http://www.ebi.ac.uk/GOA/>]
- Soares MB, Bonaldo MF, Jelene P, Su L, Lawton L, Efstratiadis A: **Construction and characterization of a normalized cDNA library**. *Proc Natl Acad Sci U S A* 1994, **91**:9228-9232.
- Bonaldo MF, Lennon G, Soares MB: **Normalization and subtraction: two approaches to facilitate gene discovery**. *Genome Res* 1996, **6**:791.
- Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment**. *Genome Res* 1998, **8**:175-185.
- RepeatMasker [<http://www.repeatmasker.org/>]
- Crossmatch [<http://www.genome.washington.edu/UWGC/analysis/tools/Swat.cfm>]
- Claverie JM, States D: **Information enhancement methods for large scale sequence analysis**. *Computers Chem* 1993, **17**:191-201.
- Blast at NCBI [<http://www.ncbi.nlm.nih.gov/BLAST/>]
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215**:403-410.
- Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J: **TIGR gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets**. *Bioinformatics* 2003, **19**:651-652.
- University of Delaware EST collection [<http://www.chick.est.udel.edu/>]
- GenomicSolutions [<http://www.genomicsolutions.com/showPage.php?title=GeneMachines%20HybStation>]
- UCSC genome bioinformatics site [<http://genome.ucsc.edu/>]
- TIGR gene indices [<http://www.tigr.org/tdb/tgi/>]
- Quackenbush J, Cho J, Lee D, Liang F, Holt I, Karamycheva S, Parvizi B, Pertea G, Sultana R, White J: **The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species**. *Nucleic Acids Res* 2001, **29**:159-64.
- Gene Ontology Consortium: **Gene Ontology: tool for the unification of biology**. *Nat Genet* 2000, **25**:25-29.